DOCUMENT RESUME

ED 209 350 TM 810 909

AUTHOR Schratz, Mary K.; Wellens, Barrie
TITLE Minority Panel Review in the Development of an Achievement Test.
INSTITUTION Psychological Corp., New York, N.Y.
PUB DATE 81
NOTE 9p.; Paper presented at the Annual Meeting of the American Psychological Association (Los Angeles, CA, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Advisory Committees; Ethnic Bias; *Item Analysis; Language Usage; *Minority Groups; Racial Bias; Sex Bias; *Sex Stereotypes; Social Bias; *Test Bias; Test Construction
IDENTIFIERS Stanford Achievement Tests; Stanford Early School Achievement Test; Stanford Test of Academic Skills

ABSTRACT

This study describes the role of an advisory panel of minority educators in the development of the seventh edition of Stanford Achievement Test (SAT) and the second editions of Stanford Early School Achievement Test (SESAT) and Stanford Test of Academic Skills (TASK). Detailed information on the formation of this panel and the review process in which it has participated in the development of SESAT/SAT/TASK is presented. Eight categories dealing with sexual bias are outlined, as are three of ethnic stereotyping. Also, regional biasing (especially vocabulary) was identified and methods of solving it shown. Figures given as the result of the panel indicate that there is little bias in the item analysis edition; bias is detected to a greater degree in reading passages and in social science items. Bias sensitivity on the part of the test developers may minimize bias, but it does not always eliminate it, particularly in its more subtle forms. Minority group educators may be best equipped for detecting subtle bias. (CE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED209350

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

MINORITY PANEL REVIEW IN THE DEVELOPMENT OF AN ACHIEVEMENT TEST

Mary K. Schratz
Barrie Wellens
The Psychological Corporation

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. K. Schratz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 609

Paper presented at the Annual Meeting of the American Psychological Association, Los Angeles, 1981.

The quality of any measuring instrument is dependent upon the expertise exercised in its construction. Potential bias against minority groups in standardized tests is an issue of concern not only to the educational community, but to the public sector as well. For this reason, an item review process by minority group educators has been implemented at several stages in the development of the seventh edition of Stanford Achievement Test (SAT) and the second editions of Stanford Early School Achievement Test (SESAT) and Stanford Test of Academic Skills (TASK), to be published in 1982. This paper provides a description of the formation of the Stanford Advisory Panel and the review process in which it has participated in the development of the series.

I. Panel Formation and Orientation

Panel members were selected on the basis of recommendations made by authors, marketing representatives, and staff of The Psychological Corporation. The Panel members are prominent members of the educational community. Ten minority group members comprise the Stanford Advisory Panel. Three are black, four are Hispanic, two represent women, and one is an American Indian. The Panel members were brought to New York City, prior to construction of the item analysis edition, to meet as a group with the authors and publisher in order to exchange views. The test authors described in detail the history and philosophy of the battery, and the construction and rationale of individual subtests. Bias was discussed and defined by the entire group in order to orient the Panel to its task. Purpose, plans, and procedures for item review were then carefully explained and simulated in a workshop setting. Examples of test items not included in the Stanford battery, which might be considered biased, were examined. Examples of some of the many ways in which sex bias may emerge were provided, along with possible solutions. These were of the following nature:

(1) Those using the terms "man", or "mankind", to represent humanity or using "he", or "him," as general referents.

Example		Possible Solution	
Someone who has run away from his duty is a --		Soldiers who run away from their responsibilities are known as --	
sprinter	fledgling	sprinters	fledglings
*deserter	hero	*deserters	heros

(2) Those depicting members of either sex in stereotyped activities, such as women cooking or sewing; men working outside the home; boys playing sports; and girls playing with dolls.

Example	Possible Solution
Sam spends a lot of time on water sports. He and his dad often go --	Sam spends a lot of time on water sports. He and his mother often go --
*sailing bowling hunting skating	*sailing bowling hunting skating
on a nearby --	on a nearby --
pool field train *lake	pool field train *lake

(3) Those depicting members of either sex as experiencing stereotyped emotions, such as girls fearful or crying and boys brave or angry.

Example	Possible Solution
Susan was <u>afraid</u> of the snake.	Susan and Don were <u>afraid</u> to cross the street.

(4) Those depicting either sex in stereotyped occupations, such as men as doctors, lawyers, or plumbers and women as nurses or secretaries.

Example	Possible Solution
(Picture of a man building a house)	Change content or balance by using a man in a non-traditional occupation in another item
baker butcher *builder calendar captain *carpenter content *construction correction	

(5) Those portraying either sex as having stereotyped characteristics, such as men being strong, brave, or scientific, while women are seen as flighty or pretty.

Example	Possible Solution
(picture of a male lifeguard saving a child)	Change lifeguard to a woman or balance with a man showing other characteristics
*help carpet humor *save serve shiver review refuse *rescue	

(6) Those portraying either sex in stereotyped situations, such as boys as creative and successful, and girls as needing help to solve problems.

Example	Possible Solution
A listening passage showing a boy making a glider and demonstrating it to a girl.	Change sex of characters and/ or balance with other content.

(7) Those using demeaning labels.

Example	Possible Solution
woman driver	driver
ladies and men	women and men
the fair sex	women

(8) Those showing art bias, such as a lack of representation of both sexes, a lack of mixed-sex groups and activities, and inappropriate dress (for example, girls always in dresses and ribbons).

Examples of some of the many ways in which ethnic or class bias may emerge were also provided. These examples were of the following nature:

(1) Those representing demeaning or offensive material.

Example	Possible Solution
A sentence reading item in which an Indian is rubbing two sticks together to make a fire.	Eliminate the item.

(2) Those in which bias of a more subtle type may emerge.

Example	Possible Solution
A reading passage in which two boys are putting a new roof on a house.	Eliminate the passage for bias against large-city children.

(3) Those showing art bias, such as lack of minority group representation, or lack of variety within minority groups.

(old artwork)

(revised artwork)

- The teacher put three under two.
- The three is beside the two.
- The teacher writes a letter.

Last, examples of ways in which regional bias may emerge were provided. Examples included those in which: (1) terms are not commonly used nationwide, e.g., "brown paper bag" where the term "sack" or "tote" is used in other parts of the country; and (2) terms have different connotations in different parts of the country.

With this orientation to the task, Panel members were then asked to indicate their preferences in regard to the particular Stanford subtests and levels they would like to review. Indicated preferences were recorded and, based on those preferences with consideration given to achieving race and sex balance, Panel members were assigned to one of two groups. The first group of reviewers was responsible for reviewing the lower test levels of the item analysis edition; the second group was responsible for reviewing the upper levels. The review task was thus split between the two groups. For each Stanford subtest, at each test level, the minimum number of reviewers necessary was determined. (For example, it was decided that more reviewers should examine the Reading and Language subtests than Mathematics Computation, which consists of numerical operations.) Panel members were then assigned particular subtests based on their indicated preferences, their educational backgrounds, and special need to have a particular group represented, e.g., an American Indian person must review Social Science. An effort was made to divide test materials among Panel members equally. At a given test level, no two members of the same racial group were assigned the same subtest for review. In addition, for both the lower and the upper level groups, an effort was made to assign a particular subtest across all lower or upper levels to the same individual.

II. The Item Review Process

Panel members were provided with guidelines for item review, the necessary test materials, and a list of the Stanford subtests and levels for which they were responsible.

When an item appeared to be biased, the Panel was asked to paste a flag on the test booklet. On the flag, they were to code the type of bias present, referring to the following list of coded Types of Content Bias:

TYPES OF CONTENT BIAS*

A. Status

Are the members of a particular group shown in situations that do not involve authority or leadership?

B. Stereotype

Are the members of a particular group portrayed as uniformly having certain: (1) aptitudes, (2) interests, (3) occupations, or (4) personality characteristics?

C. Familiarity

Is there greater opportunity on the part of members of one group to (1) be acquainted with the vocabulary, or (2) experience the situation or become acquainted with the processes presented by an item?

D. Offensive Use of Language

(1) Has a demeaning label been applied, or (2) has a male term been used where a neutral term could be substituted?

E. Other

Please be specific.

*Some of these categories of content bias represent modifications of those presented by R.B. Ekstrom in a paper entitled "Issues of Test Bias and Validity", presented at the Annual Meeting of the American Psychological Association, New York City, 1979.

Panel members were to indicate any suggestions they had for improving the "biased" item. In addition, they were instructed not to attempt to count the frequency of appearance of members of a particular minority or sex group, since balance, in this respect, would be achieved in the final edition of the battery to which they would also have the opportunity to react. (The item analysis edition consisted of four test forms per level from which two final forms would be assembled.)

III. Findings

Overall, reviewers found little evidence of bias in the Stanford items comprising the item analysis edition. Yet, bias was detected to a greater degree in some content areas than in others. Bias was not detected in any of the 240 items constituting the Primary 1 through Advanced level Mathematics Concepts test. Some Reading Comprehension and Social Science items did, however, contain evidence of bias. At each test level, bias was detected, on the average, in less than 10 percent of the Reading passages and Social Science items. These findings seem to suggest that: (1) bias is more likely to be present in certain content areas; and (2) while bias sensitivity on the part of test developers may minimize bias, it does not always eliminate it, particularly in its more subtle forms.

Further examination of those Reading Comprehension passages and Social Science items designated by reviewers as biased revealed that certain types of bias emerged more frequently than others. Of the total number of occurrences of passages or items designated as biased, 45 percent exhibited familiarity bias, 25 percent showed stereotype bias, 15 percent were detected as showing offensive use of language, and 15 percent exhibited other types of bias.

Reading passages designated as showing familiarity bias were, for the most part, those perceived as removed from the experience of the lower-class urban child. Reading passages and Social Science items in which stereotype bias emerged tended to be those in which minority group members were seen as having certain stereotyped personality characteristics or occupations. (A Reading Comprehension passage was found to be biased, for example, in that it describes a woman who is upset because she is late in meeting a train). Reviewers were also sensitive to what they considered to be offensive use of language in Reading passages or in Social Science items. (For example, the stem of a Social Science item was: "A value about land held by many Indians was that..." An American Indian reviewer suggested changing "was" to "is," in that Indians still value their land.) Other types of bias were also detected by reviewers. For instance, it was found that some items included generalities that misrepresented the interests of minority groups. (A Social Science item stating that federal land reserves are usually created in the interest of conservation was considered biased, for example, since Indian reservations are also federal reserves.)

IV. The Task-at-Hand

Each comment or issue raised by a reviewer was carefully considered in the selection of items for the final edition of this test battery. Panel members are currently having a final look at the manuscript prior to standardization. The reviewers have been asked to point out anything that strikes them as biased; they have also been asked to give their general reactions to the test materials. All reviews submitted so far are very positive.

V. Conclusions

Reviewers found little evidence of bias in the item analysis edition of Stanford Achievement Test. Bias sensitivity on the part of test developers appeared to minimize its effect; yet, bias did emerge, more frequently in some content areas than others and often in subtle ways. Certain types of content bias were present in greater frequency than others. In 45 percent of the occurrences of bias, reviewers found that there was greater opportunity on the part of members of one group to (1) be acquainted with the vocabulary; or (2) experience the situation or become acquainted with the processes presented by an item. This finding suggests that minority group educators may be best equipped for the detection of bias of this type.